

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/11776626>

# Linear prediction coding analysis and self-organizing feature map as tools to classify stress calls of domestic pigs (*Sus scrofa*)

Article in *The Journal of the Acoustical Society of America* · October 2001

DOI: 10.1121/1.1388003 · Source: PubMed

CITATIONS

52

READS

150

3 authors, including:



**Birger Puppe**

Leibniz Institute for Farm Animal Biology

151 PUBLICATIONS 2,715 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Functional lateralisation in domestic pigs (*Sus scrofa*) [View project](#)



Impulse control in pigs [View project](#)

# Linear prediction coding analysis and self-organizing feature map as tools to classify stress calls of domestic pigs (*Sus scrofa*)

Peter-Christian Schön,<sup>a)</sup> Birger Puppe, and Gerhard Manteuffel

Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere, Forschungsbereich Verhaltensphysiologie, Wilhelm-Stahl-Allee 2, D-18196 Dummerstorf, Germany

(Received 15 August 2000; revised 23 April 2001; accepted 1 June 2001)

It is assumed that calls may give information about the inner (emotional) state of an animal. Hence, in the last years sound analysis has become an increasingly important tool for the interpretation of the behavior, the health condition, and the well-being of animals. A procedure was developed that allows the characterization, classification, and visualization of the cluster structures of stress calls of domestic pigs (*Sus scrofa*). Based on the acoustic model of the sound production the extraction of features from calls was performed with linear prediction coding (LPC). A vector-based self-organizing neuronal network was trained with the determined LPC coefficients, resulting in a feature map. The cluster structure of the calls was then visualized with a unified matrix and the neurons were labeled for their input origin. The basic applicability of the procedure was tested by using two examples which were of special interest for a possible evaluation of the normal farming practice. The procedure worked well both in discriminating individual piglets by their scream characteristics and in classifying pig stress calls vs other calls and noise occurring under normal farming conditions. © 2001 Acoustical Society of America. [DOI: 10.1121/1.1388003]

PACS numbers: 43.60.Lq, 43.80.Ev, 43.60.Qv [WA]

## I. INTRODUCTION

Vocalization may provide a useful tool for evaluating the emotional state of animals under captive and natural conditions (Jürgens, 1979; Crowell Comuzzi, 1993; Mulligan *et al.*, 1994; Weary and Fraser, 1995a, b; Schrader and Todt, 1998). The important advantage of this approach is a relatively objective, noninvasive, and real-time monitoring of emotions related to environmental changes.

Husbandry may cause stress in pigs by various reasons. The stressors activate the hypothalamo–pituitary–adrenal (HPA) and the sympathico–adrenomedullary (SAM) axes via the brain's sensory and limbic pathways. These pathways also reach central motor centers which eventually trigger behavioral stress responses. One such response is vocalization being performed by sets of muscles which are located around the pulmonary–pharyngeal tract.

In pigs, the outcome is a rather sustained cry with high-frequency bands that may be highly dynamical. It is well-known from farming practice that some handling procedures, especially the restraint of animals, can induce a number of vocalizations that may reflect discomfort or distress. There is evidence that peripheral endocrine stress responses are accompanied by changing rates of specific types of vocalizations (Schrader and Rohn, 1997; Schrader and Todt, 1998). Weary *et al.* (1998) have shown that an increased rate of high-frequency calls (>1 kHz) in young piglets is a useful indicator of the pain due to castration. The analysis and classification of pigs' screams may deliver the species' and the individual's phonetic characteristics that can be attributed to

a particular stressor. If information on this interdependence is given it will be possible to judge the individual stress perception of an animal and, thus, its state of welfare or suffering. Given suitable analytic and diagnostic tools it should be possible to recognize stress quantitatively, immediately, and noninvasively. However, such tools are not available at present. In all mammals, the problem of which features are best suited for the analysis and subsequent classification has not yet been solved.

Techniques that produce good models of an arbitrary vocalization are still missing to date because it is not easy to decide which features are relevant for the exact characterization of the call. Too few features lead to an inadequate model, too many features easily overburden the computer performance for a statistical evaluation of the data (Hammerschmidt and Todt, 1995; Schrader and Hammerschmidt, 1997; Schön *et al.*, 1999). Hence, different and adapted procedures have to be applied. If such procedures are well developed they are distinguished by a good borderline of particular sounds towards other sounds, i.e., a clear, nonambiguous classification is performed.

Because of the apparent lack of effective methods, this paper presents a procedure to discriminate, classify, and visualize vocalizations of domestic pigs. The advantage of the described approach is its ability to include the dynamic structures of the calls as well as nonlinear effects like frequency steps or bifurcations. Further, decisions on the suitability for classification of the chosen features from the calls are possible. The whole procedure results in a system that will allow an assessment (e.g., classification of stressed vs not stressed). Finally, the performance of the system is demonstrated using two examples in pigs. Whereas in the first example individual piglets were discriminated by their scream character-

<sup>a)</sup> Author to whom correspondence should be addressed; electronic mail: schoen@fhn-dummerstorf.de

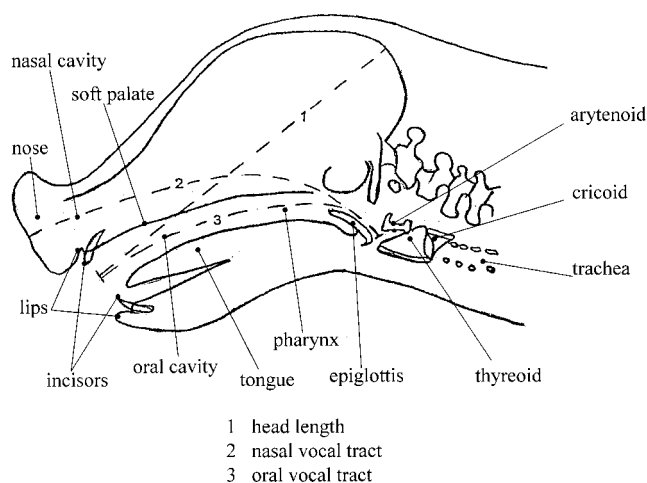


FIG. 1. Schematic drawing of the vocal tract of a piglet (redrawn from an x-ray picture).

istics, the second example shows the ability of the system to classify unknown calls as stress vocalizations when compared with calls from nonstressed animals or noises which occurred under normal farming conditions.

## II. DESCRIPTION OF THE PROCEDURE

### A. The acoustic model of sound production

According to Fant's (1970) acoustic model of sound production, vocalizations are produced by an apparatus consisting of a power supply (the lung, the thorax, and the diaphragm), the glottis as the sound source, and the nozzle formed by the vocal-tract cavern which serves as an acoustic resonator. The articulators (nose, tongue, and the soft palate; see Fig. 1) vary the size and the diameter of the nozzle. After the source-filter model, the signal that originates from the glottis is modulated by properties of the vocal tract.

The continuous air stream from the power supply is chopped by the glottis, resulting in pressure impulses (glottal source) with a certain fundamental frequency  $f_0$  and its harmonics, which can be seen in the glottis source spectrum. The glottal source signal is then modified by the vocal tract, which is the portion of the system lying above the larynx. It includes the pharynx, the oral, and the nasal cavity. The vocal-tract shape can be varied by the specific placements of the tongue, lips, and jaw. Hence, the vocal tract operates as an all-pole linear filter that introduces resonance frequencies, the so-called formants in human speech. This procedure can be formulated mathematically as

$$X(z) = H(z) * I(z), \quad (1)$$

where  $X(z)$  is the Z transformation of the generated sound signal,  $I(z)$  the source signal of the glottis, and  $H(z)$  the digital filter of the vocal tract. Then, the Z transformation of the all-pole linear filter is

$$H(z) = \frac{1}{1 - \sum_{k=1}^q c_k z^{-k}}, \quad (2)$$

and the resulting transfer function

$$X(z) = \frac{1}{1 - \sum_{k=1}^q c_k z^{-k}} * I(z), \quad (3)$$

can be regarded as a mathematical model of the vocal tract (Rabiner and Gold, 1975; O'Shaughnessy, 1987).

### B. Linear predictive coding (LPC) as a method for feature formation and data reduction

The LPC is formed in the temporal domain from a time-sequenced data series derived from a continuous signal. It is used, for example, to extract features in the frequency domain (e.g., the frequency, amplitude, and bandwidth of resonance frequencies). In the LPC, changes in the signal are used instead of the signal itself. Thus, a sound sample  $x(n)$  out of a series is taken together with a previous sample  $x(n-1)$ . A linear prediction of the actual sample is formed as a weighted sum of the past sample. The difference (prediction error)  $e(n)$  between the two samples can be minimized by introducing  $a_1$  as a coefficient, such that

$$e(n) = x(n) - a_1 x(n-1). \quad (4)$$

The minimization of the error  $e(n)$  can best be achieved if the number of previous samples is increased, which introduces also a number of additional coefficients  $a_1, \dots, a_p$ .

$$e(n) = x(n) - a_1 x(n-1) - a_2 x(n-2) - \dots - a_p x(n-p). \quad (5)$$

Applying the Z transformation, this results in

$$E(z) = X(z) - \sum_{k=1}^p a_k z^{-k} * X(z), \quad (6)$$

or

$$X(z) = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} * E(z). \quad (7)$$

This is in formal equivalence to (3), taken  $E(z)$  as the source signal from the glottis. Thus, LPC is equivalent to the source-filter model of Fant (1970) with the predictor coefficients  $a_k$  of LPC representing the vocal-tract filter coefficients  $c_k$ . For the calculation of the  $p$  coefficients  $a_1, \dots, a_p$  we used the autocorrelation method and the Levinson-Durbin recursion.

### C. Self-organizing feature maps (SOFM)

Self-organizing feature maps (Kohonen, 1982, 1997) consist of a multiplicity of homogeneous processing units (neurons) with spatial relationships to their neighbors. In the basic setup input vectors are directly fed forward to the two-dimensional output layer (map). Using the neighborhood relationships between the neurons, the map organizes the input vectors according to their structure so that similar vectors are stored in neighboring areas of the map and different vectors in distant areas. According to the Kohonen algorithm we trained the feature map with the determined LPC coefficients (12-dimensional input vectors). The mapping of the input space can then be visualized with various methods. In the

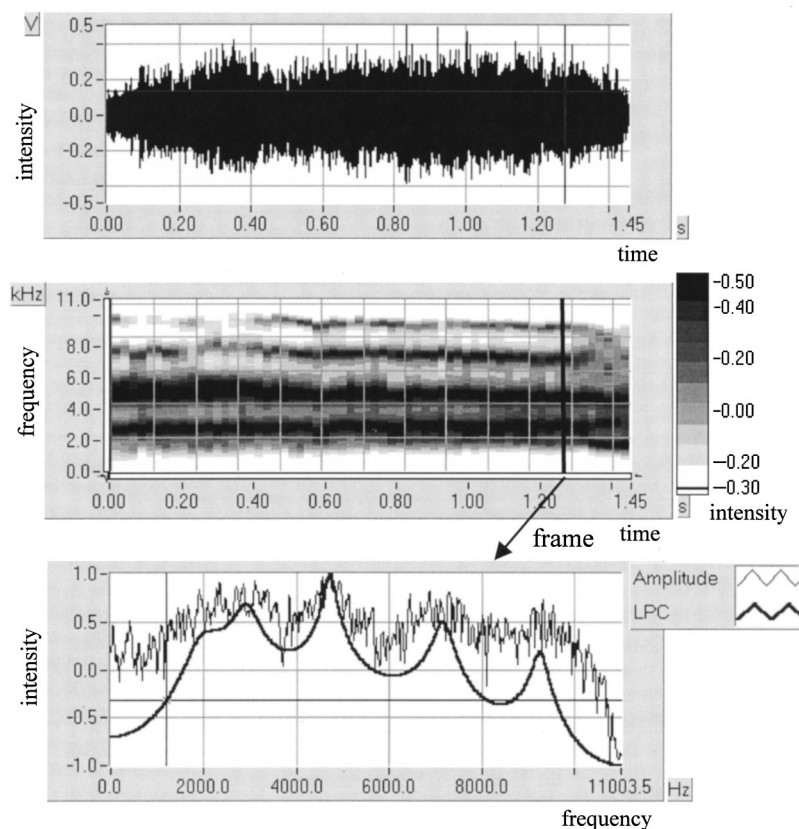


FIG. 2. Example of a stress call displayed by a piglet on the 5th day after birth. Top: Voltage signal in the time domain. Middle: Normalized logarithmic LP spectrum of the signal depicted at the top. The gray-scale density represents the intensity of the corresponding frequencies. Two resonance frequency bands at about 3 and 5 kHz are clearly visible. A narrower frequency band lies at about 8 kHz. A further, incomplete band is seen between 8 and 10 kHz. Bottom: Normalized logarithmic amplitude spectrum extracted from the FFT transformation and the LP spectrum of a single frame (46.44 ms). The local maxima of the LP spectrum coincide with the resonance spectra of the amplitude spectrum.

second step the mapping of the Kohonen network was tested with test data sets which were not included in the training data set.

A methodical problem before the self-organization can start is the choice of suitable network parameters (neurons  $X$ , neurons  $Y$ , learning radius, learning radius factor, learning rate, learning rate factor, learning steps). Still, a suitable map structure can be reached only in trying. In our case, we started with a small number of neurons ( $60 \times 60$ ) that was increased until no further improvement of the classification result and the classification parameters (maximum and mean distance—learning curve) occurred. This methodical problem could also be solved by the application of self-optimizing growing cell structures. A description of such structures is given by Fritzke (1992, 1998).

#### D. Extraction and graphical representation of the LPC coefficients

The number of the LPC coefficients  $a_1, \dots, a_p$  determines the number of the considered resonance frequencies of the vocal tract, with  $p/2$  being the number of the resonance frequencies. Using up to 18 LPC coefficients we never found more than 6 resonance frequencies in the piglets' screams. Hence, for further analyses, 12 LPC coefficients were used. In order to take into account the dynamic structure of the calls, the LPC coefficients  $a_1, \dots, a_p$  were extracted from time windows (frames) of 46.44 ms length in each case. The LP spectrum was determined from the LPC coefficients by means of a polynomial development.

Figure 2 shows the results of the spectral analyses of a stress call as amplitude and LP spectra. The dynamic struc-

ture of the call is clearly visible by the bands of the LP spectrum in the frequency domain. Some bands alter their frequency and amplitude, and the low-frequency band bifurcates at the end of the recording. In the single-amplitude spectrum (bottom), two effects of the sound production mechanism are visible. The influence of the stimulation frequency (glottal source) results in a fundamental frequency  $f_0$  at about 250 Hz and its harmonics. The harmonics are the quickly (200-Hz range) changing parts. The influence of the resonance qualities of the vocal tract causes slow changes. The LP spectrum represents a model of the vocal tract and may substitute the amplitude spectrum with respect to the resonance frequencies. Hence, it is possible to determine the resonance frequencies of the vocal tract from the local maxima of the LP spectrum.

### III. EXAMPLES FOR THE APPLICABILITY OF THE PROCEDURE

#### A. Example I: Classification of individual piglets by their scream characteristics

##### 1. Animals and recordings

The calls to be individually classified were recorded from three randomly selected piglets (*Sus scrofa*) of the "German Landrace". They were housed together with their mother and kept in a standard farrowing crate. "Stress screams" were recorded on the fifth day after birth within an interval of 2 min per animal. In this time we got between 26 and 62 screams per animal for classification. For the acoustical recordings the piglets were removed rapidly from their home pen and taken to the adjacent acoustic chamber with a minimum of disturbance. The chamber with low acoustic



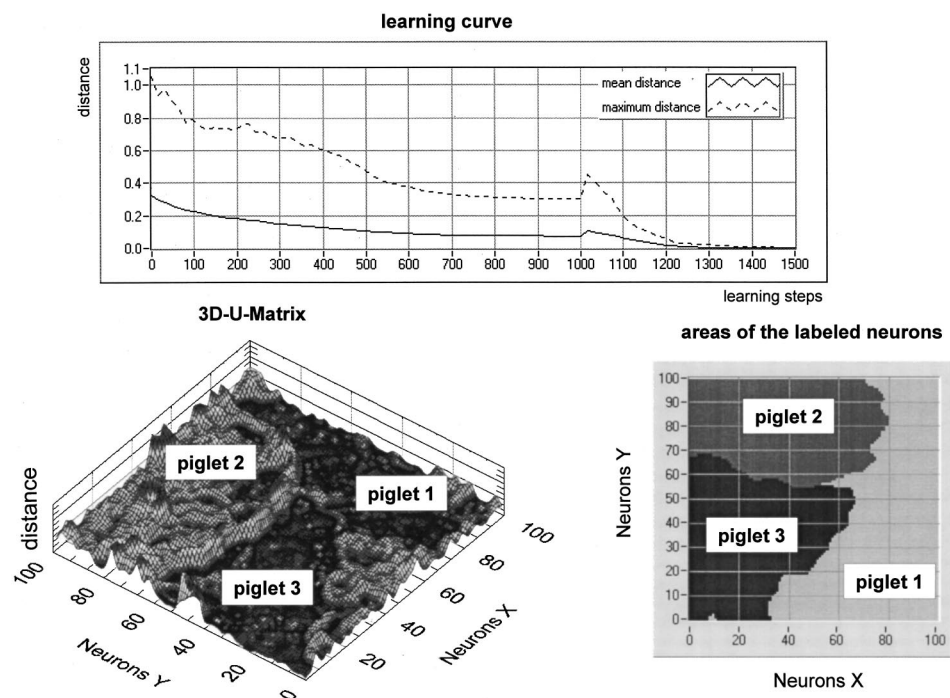


FIG. 3. Learning curve, 3D-U Matrix, and areas of the labeled neurons for the model on the 5th day after birth. The learning curve shows the quality of the training process of the SOFM with the help of the gradual descent of the mean and maximum distance of the winner neurons to the respective input vectors; the U matrix visualizes the separation of clusters by the SOFM; the areas of labeled neurons decides further about the assignment of the test vectors to the individual clusters.

reflection by the walls and the ceiling was described earlier by Schön *et al.* (1998). Screams were experimentally induced by keeping up the piglets at the thorax by a person for 1 to 2 min—a stressful situation which is quite similar to the handling immediately prior to the castration of male piglets in the normal farm animal practice, or, more common, to a situation where the sow lays down and accidentally squeezes a piglet to death. In the latter situation the screaming triggers the sow to stand up immediately (Hutson *et al.*, 1993; Weary *et al.*, 1996). All calls were recorded with a DAT recorder (Sony DCT-790) and a separate microphone (Sennheiser MKE 46). The procedures for data processing were developed and programmed using the graphical programming language LABVIEW<sup>®</sup> with the additional tool DATAENGINE V.I.<sup>®</sup> (LABVIEW<sup>®</sup>, 1998; DATAENGINE V.I.<sup>®</sup>, 1999).

## 2. Training and test phase of the system

Fifteen screams of each individual were used to train the map and further ten screams of each piglet were used to test the map. The test screams were not included in the training data set.

*a. Training phase.* The SOFM was trained with 12-dimensional input vectors derived from the 12 LPC coefficients per time window (46.44 ms). Hence, 9 to 45 LPC vectors per scream were obtained. The maps used in this experiment consisted of  $100 \times 100$  neurons and were trained in  $3 \times 500$  learning steps. The initialization of the first 1000 steps was performed with a learning radius of 60, a learning radius factor of 0.995, a learning rate of 0.999, and a learning rate factor of 0.99. The continuous training during the following 500 learning steps was performed with a learning radius of 4; all other parameters were left unchanged. Figure 3 shows the achieved learning curves that represent the gradual descent of the mean and maximum distance of the winner neurons to the respective input vectors.

In order to assess the result of the training process, a graphic representation has proven to be helpful. This can be obtained by a U-matrix (unified matrix) representation (Ultsch and Simon, 1990), where the distance between neighboring weight vectors of the map is determined and applied to the matrix. The result can be plotted three-dimensionally where the height of the hills is proportional to the distance of the neurons. In this representation information exists over the entire input space. For each neuron, the Euclidian distance to its neighbors is determined. The obtained result is depicted in Fig. 3.

The representation with the U matrix is not dependent on the dimension of the input space. Based on the topological representation of the map, the input vectors that belong to the same cluster are found in the same areas (valleys) on the map. The valleys are separated by the hills. The scatter of the single-resonance frequencies of the vocal tract over the call and the differences between the single calls of an animal are reflected by the neighborhood relationships of the neurons. The clusters 1–3 were assigned to the three examined piglets. While piglets 1 and 3 displayed a very homogeneous structure in the screams, indicating a low variability of the calls, piglet 2 delivered a strong variability within and between the screams. Nevertheless, piglet 2 was clearly separated from the other two piglets, while the separation between piglet 1 and 3 was not as strong. The U matrix makes it possible to recognize the separation of clusters by the SOFM without having exact knowledge of the input data set. This is an important aspect in sound analysis. Thus, it is possible to decide whether or not a classification result can be reached with the chosen feature vectors from the calls.

After training of the SOFM the resulting clusters were labeled using the knowledge of the individual origins of the training data set. The result is shown in Fig. 3, where the neuronal areas were labeled with respect to the three piglets.

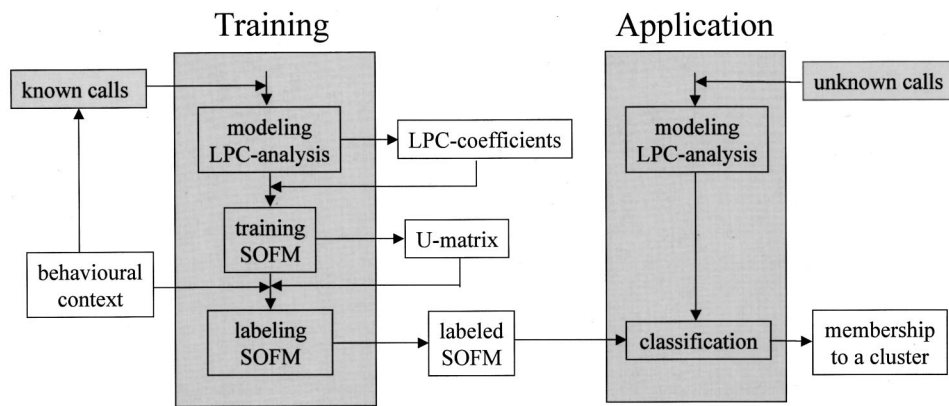


FIG. 4. Schematic representation of the process of the training and the application (testing) phase by using the SOFM.

With the labeling the training phase was finished.

*b. Testing phase.* The trained classification model was now ready for testing with unknown test vectors from screams that were not used for training (Fig. 4). The results of the assignment of the test LPC vectors are given in Table I. The misclassification rate was very low (<3%).

## B. Example II: Classification of piglet screams versus other calls and noise

### 1. Animals, calls, and recordings

The procedure is not only able to deliver an individual assignment of calls, but also a classification regarding different call types. Hence, different vocal responses of pigs regarding their environment and treatment were exemplarily recorded to be analyzed later.

First, stress screams of 19 piglets from 4 different litters (litter 1–litter 3: each with three 2-week-old piglets [training], litter 4: ten 2-week-old piglets [testing]) were recorded in the same way as described in example I. We further used the screams of 16 growing pigs (5-week-old pigs) from two different litters (litter 5: nine 5-week-old piglets [training], litter 6: seven 5-week-old piglets [testing]).

Second, grunts vocalized in various social situations were used (calls displayed in a “nonstress” context). Both the grunts of six 2-week-old piglets (three for training and three for testing) and the grunts of six 5-week-old growing pigs (three for training and three for testing) were induced by a short social isolation of the animals with vocal contact to a companion. Further, the nursing grunts (see Schön *et al.*, 1999) of ten first-lactating sows (five for training and five for testing) were involved.

Third, three examples of the background noise occurring in the housing environment of the pigs were used. They consist of arbitrary sounds (e.g., the talking of humans, air ventilation, and rattle of the equipment) without any animal calls (noise I for training, noise II for testing) or with a small rate

of animal vocalizations included in the background noise (noise III for testing). All recordings were made randomly under normal keeping conditions in the stable. The recording equipment was the same as described in example I.

### 2. Training and test phase of the system

*a. Training phase.* The map was trained with LPC vectors that were determined from screams of piglets and growing pigs. We used 7673 LPC vectors to characterize the stress (scream) area of the Kohonen map (screams from litter 1–3 and litter 5). On the other side we used LPC vectors from grunts of three piglets (724 LPC vectors), three growing pigs (2716 LPC vectors), and five first-lactating sows (178 LPC vectors), and noise I (1263 LPC vectors). Hence, the non-stress area of the Kohonen map was characterized by 4881 LPC vectors and the map was trained with a total of 12 554 LPC vectors. The maps used in this experiment consisted of  $150 \times 150$  neurons and were trained in  $3 \times 500$  learning steps. The initialization of the first 1000 steps was performed with a learning radius of 80, a learning radius factor of 0.995, a learning rate of 0.999, and a learning rate factor of 0.99. The continuous training during the following 500 learning steps was performed with a learning radius of 4; all other parameters were left unchanged.

*b. Testing phase.* The labeled SOFM was then tested with unknown LPC vectors from screams (litter 4, litter 6), grunts (three 2-week-old piglets, three 5-week-old piglets, and five first-lactating sows), and noise (noise II, noise III). The findings are shown in Table II.

The results show that the classification of screams to the “stress area” was possible with a misclassification rate lower than 1%. The grunts were to >97.5 correctly attributed to the “non-stress area.” A similar classification result was achieved with noise II. In addition to noise II, noise III consisted also of a small rate of animal communication. This was a mixture of grunts, squeals, and screams (e.g. fighting for rank order, waiting and fighting in front of the trough before feeding, attempts of escaping after isolation from the social group) which might be regarded as typical examples of vocalizations in a housing environment. When a “non-stress” context is assumed (Table II) the classification shows an error of 7%.

TABLE I. Example I: Results for testing the trained SOFM with unknown stress screams of three piglets on their 5th day after birth.

| Piglet | No. of analyzed LPC vectors | Misclassification rate (%) |
|--------|-----------------------------|----------------------------|
| 1      | 380                         | 0.00                       |
| 2      | 281                         | 1.78                       |
| 3      | 272                         | 2.94                       |

TABLE II. Example II: Results for testing the trained SOFM with unknown calls or noise classified as “stress” or “nonstress” calls.

| Animals (age)          | <i>n</i> | Calls/noise          | No. of analyzed<br>LPC vectors | Type      | Misclassification<br>rate<br>means±s.d. (%) |
|------------------------|----------|----------------------|--------------------------------|-----------|---------------------------------------------|
| Piglets (2 weeks)      | 10       | screams              | 1904                           | stress    | 0.58±1.11                                   |
| Growing pigs (5 weeks) | 7        | screams              | 2476                           | stress    | 0.85±1.27                                   |
| Piglets (2 weeks)      | 3        | grunts               | 171                            | nonstress | 2.34±2.55                                   |
| Growing pigs (5 weeks) | 3        | grunts               | 245                            | nonstress | 2.04±1.70                                   |
| Sows (1st lactation)   | 5        | nursing grunts       | 60                             | nonstress | 1.67±2.89                                   |
| Noise II               | ...      | without animal calls | 1706                           | nonstress | 1.23±1.15                                   |
| Noise III              | ...      | with animal calls    | 1072                           | nonstress | 7.00±2.79                                   |

#### IV. DISCUSSION

The aim of the present paper is to introduce a procedure for the classification and identification of stress screams of domestic pigs that allows the classification of animal calls using the LPC analysis and a subsequent SOFM of the Kohonen type (Kohonen, 1992).

Previous attempts of computer-aided analysis of animal vocalization were mainly based on conventional statistical methods (e.g., a discriminant analysis) with a multiparametric approach (Schrader and Hammerschmidt, 1997; Schön *et al.*, 1999). Where features like FFT coefficients or cepstrum analysis were used in combination with neural networks, the biological meaning of the respective sounds were not in the center of interest (cf. Datum *et al.*, 1996), or the dynamic structure and/or the large amount of parameter coefficients of the sounds to be recognized demanded large and complicated systems (cf. Kohonen, 1992; Schuchardt, 1992; Reby *et al.*, 1996, 1997).

Calls of mammals often have a complex structure. In past years it was shown that the sound production is not only a deterministic process but rather the result of a highly nonlinear dynamic system (Herzel *et al.*, 1995). The model of Fant (1970) represents only an idealized view on the underlying processes. In nonhuman mammal communication the appearance of nonlinear phenomena seems to be normal and it seems to occur more often in disordered animals (Riede *et al.*, 1997, 2000; Wilden *et al.*, 1998). Thus, every kind of modeling has to take these aspects into account.

Linear predictive coding (LPC), based on a source-filter model, is compatible with the mechanism of sound generation in the vocal tract and, most importantly in our application, reduces the number of coefficients necessary for a successful classification of stress calls to 12. This is a considerable reduction compared to 512 Fourier coefficients, and still more than the 50 cepstral coefficients used in an earlier study of pig vocalization (Schön *et al.*, 1999).

The coefficients of the LPC analysis were topologically represented on the feature map. Subsequently, the U matrix delivered the number of classes on the map. The result was identical with the number of animals from which the labeled data set originated. After the training phase we tested the structure of the network with unknown test screams. The classification results clearly showed that an individual assignment of the three exemplary selected piglets was possible on the basis of their particular scream characteristics.

A further task was testing a trained network regarding the classification performance with respect to the classification of piglet screams vs other calls and noise. Here, after the training was completed, unknown stress calls from pigs had to be assigned to the trained clusters. Most of the unknown calls were correctly assigned to the area of stress screams.

In order to obtain reasonable training times and an on-line capability of the trained system, as well as a good classification, it is necessary to estimate the optimum number of LPC coefficients that are inputs to the SOFM. This, like the size of the Kohonen network, is still not solved in general. Clearly, this optimum number depends on the sounds to be discriminated and, hence, has to be evaluated with each new task. However, this method also bears some advantages. The flexible arrangement of the feature extraction as well as the arrangement of the topology of the neuronal network allows extensive simulations of the training process and makes it possible to solve a multiplicity of classification tasks and have visual access to them.

The functionality of the described approach clearly depends on the fact that pig vocalizations are more or less sustained. That means that only weak frequency modulations occur during a call. Hence, the temporal order of the LPC coefficients was irrelevant to the SOFM. However, the comparatively simple design of the LPC-SOFM procedure for stress-call classification and identification makes it possible to create an on-line monitoring system which can detect stress calls immediately, since the necessary number of LPC coefficients and the classification by the trained neural network can be calculated in real time.

Hopefully, further development of the method extending it to other types of vocalizations will enable us to create systems that are able to recognize, discriminate, and output the semantic meaning of various animal calls. Similar to the proposal of Ritter and Kohonen (1989) for the semantics of words, this could be reached by grouping the calls according to their behavioral context.

Thus, the present procedure may be used as a methodological approach to solve different analysis and classification tasks in animal vocalization. Besides the clarification of communicative aspects, it would allow us to automatically monitor behavioral responses of farm animals in a housing situation with respect to their well-being (e.g., stressed vs nonstressed) or suffering (e.g., sick vs healthy). This will be

of increasing significance in all fields of captive and restricted animals in human charge.

## ACKNOWLEDGMENTS

The authors thank the staff of the Department of Behavioral Physiology for expert technical assistance. Special thanks go to Tobias Riede for fruitful discussions and helpful comments during the work and for kindly allowing us to use one of his drawings from x rays (Fig. 1).

- Crowell Comuzzie, D. K. (1993). "Baboon vocalizations as measures of psychological well-being," *Lab. Prim. Newsletter* **32**, 5–6.
- DATAENGINE VI.® (1999). *User Manual, Function Reference, Tutorials, Basics* (MIT-Management Intelligenter Technologien GmbH, Aachen, Germany).
- Datum, M. S., Palmieri, F., and Moiseff, A. (1996). "An artificial neural network for sound localization using binaural cues," *J. Acoust. Soc. Am.* **100**, 372–383.
- Fant, G. (1970). *Acoustic Theory of Speech Production* (2nd printing, Mouton, The Hague).
- Fritzke, B. (1992). "Wachsende Zellstrukturen—ein selbstorganisierendes neuronales Netzwerk," Ph.D. thesis (unpublished), Universität Erlangen-Nürnberg.
- Fritzke, B. (1998). *Vektorbasierte Neuronale Netze* (Shaker GmbH).
- Hammerschmidt, K., and Todt, D. (1995). "Individual differences in vocalizations of young Barbary macaques (*Macaca sylvanus*): A multi-parametric analysis to identify critical cues in acoustic signaling," *Behaviour* **132**, 381–399.
- Hutson, G. D., Price, E. O., and Dickenson, L. G. (1993). "The effect of playback volume and duration on the response of sows to piglet distress calls," *Appl. Anim. Behav. Sci.* **37**, 31–37.
- Herzel, H., Berry, B., Titze, I., and Steinecke, I. (1995). "Nonlinear dynamics of the voice: Signal analysis and biomechanical modeling," *Chaos* **5**, 1–5.
- Jürgens, U. (1979). "Vocalization as an emotional indicator. A neuroethological study in the squirrel monkey," *Behaviour* **69**, 88–117.
- Kohonen, T. (1982). "Self-organized formation of topologically correct feature maps," *Biol. Cybern.* **43**, 59–69.
- Kohonen, T. (1992). "How to make a machine transcribe speech," in *Application of Neural Networks*, edited by H. G. Schuster (Chemie, Weinheim), pp. 25–34.
- Kohonen, T. (1997). *Self-organizing Maps*. Springer Series in Information Sciences, Vol. 30, 2nd ed. (Springer, Berlin, Heidelberg, New York).
- LABVIEW® (1998). *Complete Software Documentation* (National Instruments Corporation, Austin, TX).
- Mulligan, B. E., Baker, S. C., and Murphy, M. R. (1994). "Vocalizations as indicators of emotional stress and psychological well being in animals," *Anim. Welfare Inform. Center Newsletter* **5**, 3–4.
- O'Shaughnessy, D. (1987). *Speech Communication* (Addison-Wesley, Reading, MA).
- Rabiner, L. R., and Gold, B. (1975). *Theory and Application of Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).
- Reby, D., Joachim, J., Lauga, J., Cargnelli, B., and Gonzalez, G. (1996). "Using voice recognition as a tool in population biology and management," 1st International Conference on Methods and Techniques in Behavioural Research, Utrecht, 16–18 October 1996.
- Reby, D., Lek, S., Dimopoulos, I., Joachim, J., Lauga, J., and Aulagnier, S. (1997). "Artificial neural networks as classification method in the behavioural sciences," *Behav. Processes* **40**, 35–43.
- Riede, T., Tembrock, G., Herzel, H., and Brunnberg, L. (1997). "Vocalization as indicator for disorders in mammals," 137th meeting of the Acoustical Society of America, San Diego, 1–5 December 1997.
- Riede, T. (2000). "Vocal changes in animals during disorders," Ph.D. thesis (unpublished), Humboldt University, Berlin.
- Ritter, H., and Kohonen, T. (1989). "Self-organizing semantic maps," *Biol. Cybern.* **61**, 241–254.
- Schrader, L., and Hammerschmidt, K. (1997). "Computer-aided analysis of acoustic parameters in animal vocalization: A multi-parametric approach," *Bioacoustics* **7**, 247–265.
- Schrader, L., and Rohn, C. (1997). "Lautäußerungen von Hausschweinen als Indikator für Stressreaktionen," *Landbauforsch. Völkenrode* **47**, 89–95.
- Schrader, L., and Todt, D. (1998). "Vocal quality is correlated with levels of stress hormones in domestic pigs," *Ethology* **104**, 859–876.
- Schön, P. C., Puppe, B., and Manteuffel, G. (1998). "A sound analysis system based on LABVIEW® applied to the analysis of suckling grunts of domestic pigs (*Sus scrofa*)," *Bioacoustics* **9**, 119–133.
- Schön, P. C., Puppe, B., Gromyko, T., and Manteuffel, G. (1999). "Common features and individual differences in nurse grunting of domestic pigs (*Sus scrofa*): A multi-parametric analysis," *Behaviour* **136**, 49–66.
- Schuchardt, J., Gruel, J. C., Luthje, N., Molgedy, L., Radons, G., and Schuster, H. G. (1992). "Neural networks for the classification of sound patterns," in *Application of Neural Networks*, edited by H. G. Schuster (Chemie, Weinheim), pp. 239–249.
- Ultsch, A., and Simon, H. P. (1990). "Kohonen's self-organizing feature map for exploratory data analysis," in *Proceedings of International Neural Networks* (Kluwer Academic, Paris), pp. 305–308.
- Weary, D. M., and Fraser, D. (1995a). "Calling by domestic piglets: Reliable signals of need?" *Anim. Behav.* **50**, 1047–1055.
- Weary, D. M., and Fraser, D. (1995b). "Signaling need: Costly signals and animals welfare assessment," *Appl. Anim. Behav. Sci.* **44**, 159–169.
- Weary, D. M., Pajor, E. A., Thompson, B. K., and Fraser, D. (1996). "Risky behavior by piglets: A trade off between feeding and risk of mortality by maternal crushing?" *Anim. Behav.* **51**, 619–624.
- Weary, D. M., Braithwaite, L. A., and Fraser, D. (1998). "Vocal response to pain in piglets," *Appl. Anim. Behav. Sci.* **56**, 161–172.
- Wilden, I., Herzel, H., Peters, G., and Tembrock, G. (1998). "Subharmonics, biophonation, and deterministic chaos in mammal vocalization," *Bioacoustics* **9**, 171–196.